Integrative Molecular Concepts Analysis of Prostate Cancer Progression

Scott A. Tomlins[1], Rohit Mehra[1,6], Daniel R. Rhodes[1,2,6], Xuhong Cao[1], Lei Wang[1], John T. Wei[3,4,6], Mark A. Rubin[7,8], Kenneth J. Pienta[3,4,5,6], Rajal B. Shah[1,3,4,6], Arul M. Chinnaiyan[1,2,3,4,6]

Departments of Pathology[1], Bioinformatics[2], Urology[3], Michigan Urology Center[4], Internal Medicine[5], Comprehensive Cancer Center[6], University of Michigan Medical School, Ann Arbor, MI 48109

Department of Pathology[7], Brigham and Women's Hospital, Dana-Farber Cancer Institute[8], Harvard Medical School, Boston, MA

Abstract

Similar to other epithelial cancers, prostate cancer progression has been defined histologically as a transition from benign epithelium, to precursor lesions such as high grade prostatic intraepithelial neoplasia (PIN), to adenocarcinoma, and ultimately to metastatic disease. Despite efforts to profile prostate cancer progression using DNA microarrays, the genetic alterations and biological processes that correlate with the observed histological progression are unclear. Using laser capture microdissection to isolate over 100 specific cell populations, we report the profiling of prostate cancer progression from benign epithelium to metastatic disease. By analyzing these expression signatures in the context of 15,000 "molecular concepts", or sets of biologically related genes, we generated a model of prostate cancer progression. Molecular concepts that demarcate critical transitions in prostate cancer progression include protein biosynthesis, ETS family transcriptional targets, androgen signaling, and cell proliferation. Of note, high grade prostate cancer (Gleason Pattern 4) exhibits an attenuated androgen signature relative to low grade prostate cancer (Gleason Pattern 3). Taken together, we demonstrate that analyzing gene expression signatures in the context of a compendium of molecular concepts has utility in understanding disease biology.

Methods

Laser Capture Microdissection

Laser Capture Microdissection (LCM) was performed from frozen tissue sections with the SL Microtest device using µCUT software (MMI). Approximately 10,000 cells were captured for each sample.

RNA Amplification and Hybridization

Exponential RNA amplification was performed using a TransPlex Whole Transcriptome Amplification (WTA) kit (Rubicon Genomics, Ann Arbor, MI) as described (Tomlins et al. 2006, Neoplasia 8:153-162). Amplified cDNA was hybridized to 20K element cDNA microarrays.

Data Anaylsis

The data set was loaded into the Oncomine database (www.oncomine.org) for identification of gene signatures and automated gene set enrichment analysis against all concepts in the Molecular Concepts Map.

Acknowledgements

This work was supported by the Early Detection Research Network (EDRN) and the Specialized Program of Research Excellence (SPORE)

Conclusions

- Combining laser capture microdissection with exponential RNA amplification allows for specific profiling
- Proliferation related concepts are over-expressed in PCA progression, particularly from localized to metastatic PCA
- Androgen signaling activity defines several key transitions in PCA progression, with increased activity from benign to PIN, and decreased activity in PCA progression, localized to metastatic disease, and importantly, low to high Gleason grade PCA
- Increased protein biosynthesis concepts define the benign to PIN transition, consistent with an enlarged nucleolus being the defining histological feature of PIN
- Changes in protein biosynthesis and ETS transcriptional targets parallel changes in androgen signaling activity, suggesting a functional link
- Marked over-expression of ETS transcription factors through gene fusions with TMPRSS2 likely defines the PIN to PCA transition in a majority of cases
- Integrating expression profiling with a compendium of molecular concepts is useful for understanding disease biology, confirming previous hypotheses and generated novel genes and concepts involved in PCA progression

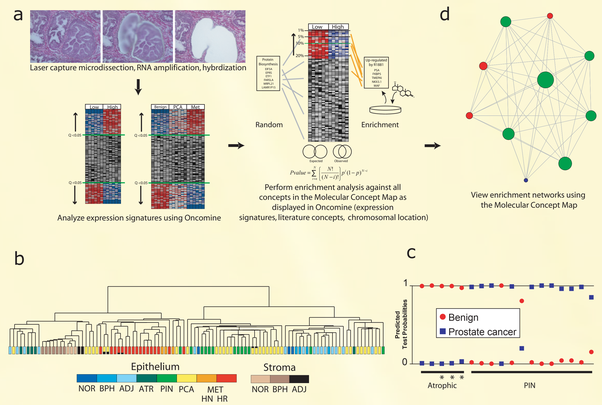

Figure 1. Integrative analysis of molecular concepts in prostate cancer progression. a. Schematic of the experimental approach. **b**. Unsupervised hierarchical clustering of the complete data set. Samples are colored by class according to the color scale. **c**. Prediction Analysis of Microarrays (PAM) demonstrates that while PIN and PCA share similar gene expression signatures, atrophic samples are more similar to benign prostate epithelium.

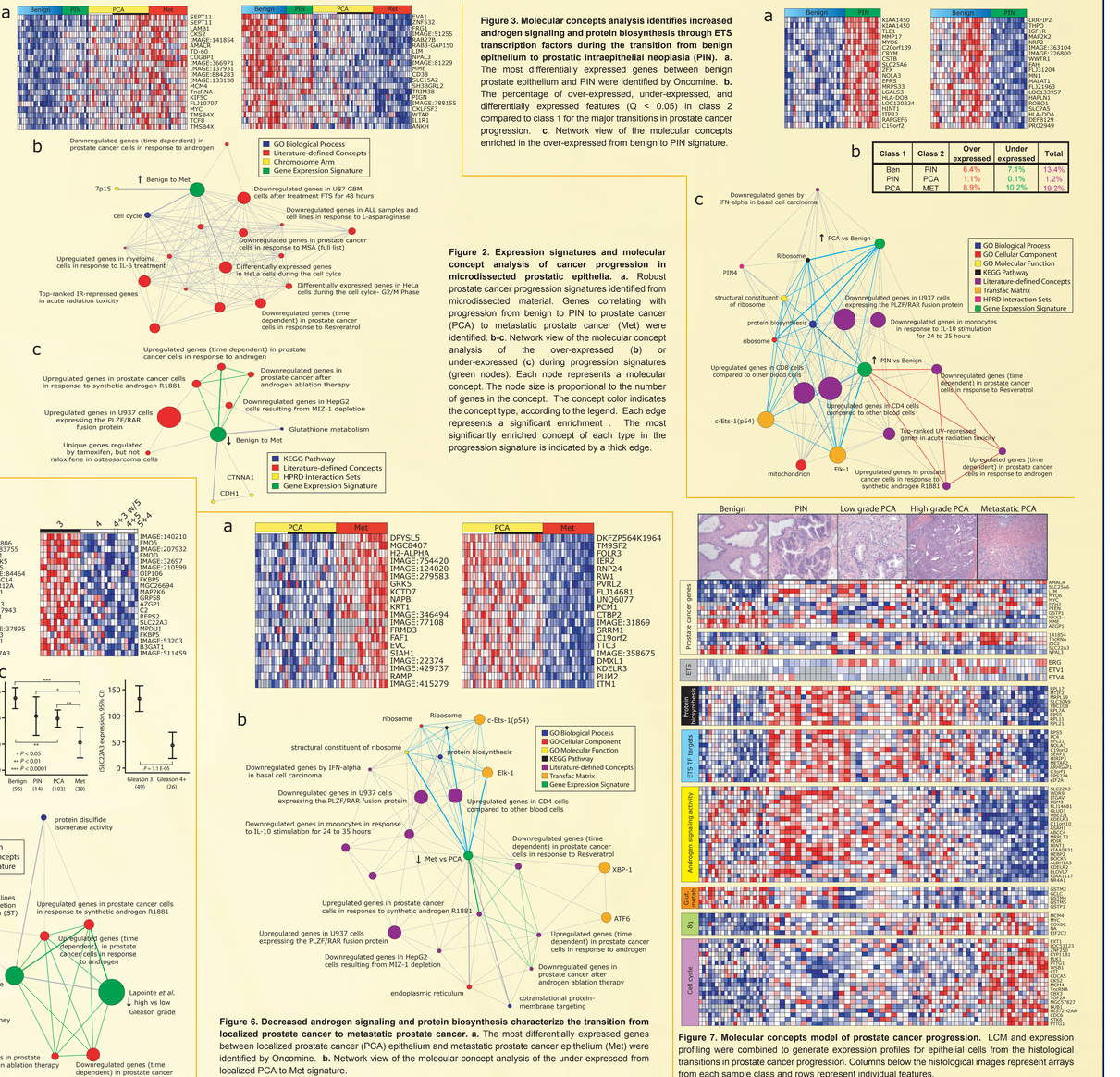

Figure 2. Expression signatures and molecular concept analysis of cancer progression in microdissected prostatic epithelia. a. Robust prostate cancer progression signatures identified from microdissected material. Genes correlating with progression from benign to PIN to prostate cancer (PCA) to metastatic prostate cancer (Met) were identified. **b-c**. Network view of the molecular concept analysis of the over-expressed (**b**) or under-expressed (**c**) during progression signatures (green nodes). Each node represents a molecular concept. The node size is proportional to the number of genes in the concept. The concept color indicates the concept type, according to the legend. Each edge represents a significant enrichment . The most significantly enriched concept of each type in the progression signature is indicated by a thick edge.

Figure 3. Molecular concepts analysis identifies increased androgen signaling and protein biosynthesis through ETS transcription factors during the transition from benign epithelium to prostatic intraepithelial neoplasia (PIN). a. The most differentially expressed genes between benign prostate epithelium and PIN were identified by Oncomine. **b.** The percentage of over-expressed, under-expressed, and differentially expressed features (Q < 0.05) in class 2 compared to class 1 for the major transitions in prostate cancer progression. **c**. Network view of the molecular concepts enriched in the over-expressed from benign to PIN signature.

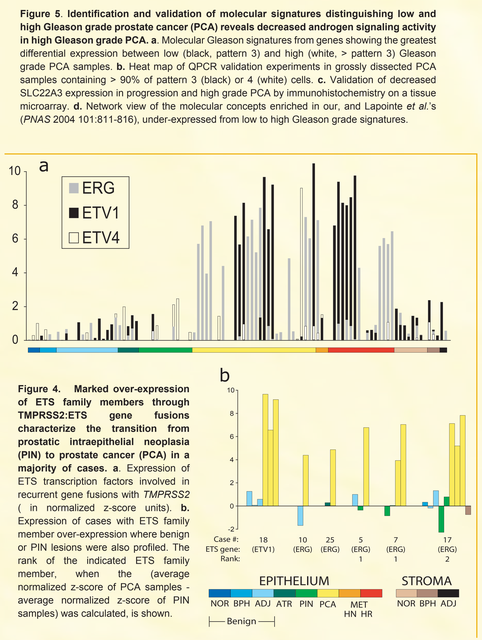

Figure 5. Identification and validation of molecular signatures distinguishing low and high Gleason grade prostate cancer (PCA) reveals decreased androgen signaling activity in high Gleason grade PCA. a. Molecular Gleason signatures from genes showing the greatest differential expression between low (black, pattern 3) and high (white, > pattern 3) Gleason grade PCA samples. **b.** Heat map of QPCR validation experiments in grossly dissected PCA samples containing > 90% of pattern 3 (black) or 4 (white) cells. **c.** Validation of decreased SLC22A3 expression in progression and high grade PCA by immunohistochemistry on a tissue microarray. **d.** Network view of the molecular concepts enriched in our, and Lapointe *et al.*'s (*PNAS* 2004 101:811-816), under-expressed from low to high Gleason grade signatures.

Figure 4. Marked over-expression of ETS family members through TMPRSS2:ETS gene fusions characterize the transition from prostatic intraepithelial neoplasia (PIN) to prostate cancer (PCA) in a majority of cases. a. Expression of ETS transcription factors involved in recurrent gene fusions with *TMPRSS2* (in normalized z-score units). **b.** Expression of cases with ETS family member over-expression where benign or PIN lesions were also profiled. The rank of the indicated ETS family member, when the (average normalized z-score of PCA samples - average normalized z-score of PIN samples) was calculated, is shown.

Figure 6. Decreased androgen signaling and protein biosynthesis characterize the transition from localized prostate cancer to metastatic prostate cancer. a. The most differentially expressed genes between localized prostate cancer (PCA) epithelium and metastatic prostate cancer epithelium (Met) were identified by Oncomine. **b.** Network view of the molecular concept analysis of the under-expressed from localized PCA to Met signature.

Figure 7. Molecular concepts model of prostate cancer progression. LCM and expression profiling were combined to generate expression profiles for epithelial cells from the histological transitions in prostate cancer progression. Columns below the histological images represent arrays from each sample class and rows represent individual features.